UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/797,992	03/11/2004	Eric R. Buhrke	CML01339T	5169

22917 7590 11/17/2006

MOTOROLA, INC.
1303 EAST ALGONQUIN ROAD
IL01/3RD
SCHAUMBURG, IL 60196

EXAMINER	
STORM, DONALD L	
ART UNIT	**PAPER NUMBER**
2626	

DATE MAILED: 11/17/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No. 10/797,992	Applicant(s) BUHRKE, ERIC R.
	Examiner Donald L. Storm	Art Unit 2626

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on *17 October 2006*.

2a)☒ This action is **FINAL**. 2b)☐ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1,2,4,5 and 7-15* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1,2,4,5 and 7-15* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☐ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .
5) ☐ Notice of Informal Patent Application
6) ☐ Other: _____.

DETAILED ACTION

1. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

Claim Informalities

2. Claim 2 is objected to under 37 CFR 1.75(c), as being of improper dependent form for failing to further limit the subject matter of a previous claim, because the 10-msec latency is inherently less than 100-msec latency. The Applicant is required to cancel the claim, or amend the claim to place the claim in proper dependent form, or rewrite the claim in independent form.

3. Claims 1, 4, 12, 13, 14, and by dependency claims 2, 4-5, 7-11 and 15, are objected to under 37 CFR 1.75(a) because some phrases need clarification of potentially confusing informalities, as follows.

a. In claim 1 (line beginning *positive*), should the word "integers" be --integer--?

b. In claim 1 (line beginning *classification*), the phrase ", wherein N is" should not be repeated here. The symbol N has previously been defined and should not be redefined to another definition. Repeating the same definition as a duplicate limitation may be confusing.

c. In claim 4 (line 2), should the word "visemes" (second occurrence) be --viseme--?

d. Claim 12 contains the same (two) informalities as claim 1.

e. Claim 13 contains the same (two) informalities as claim 1.

f. In claim 14 (line beginning *frame classification*), the phrase ", wherein N is" should not be repeated here. The symbol N has previously been defined and should not be redefined to another definition. Repeating the same definition as a duplicate limitation may be confusing.

Claim Rejections - 35 USC § 112

4. The following is a quotation of the second paragraph of 35 U.S.C. 112:

> The specification shall conclude with one or more claims particularly pointing out and distinctly claiming the subject matter which the applicant regards as his invention.

5. Claims 7-8 are rejected under 35 U.S.C. 112, second paragraph, as being indefinite for failing to particularly point out and distinctly claim the subject matter which applicant regards as the invention.

6. Claim 7, and by dependency claim 8, are indefinite because they are incomplete. Dependency from claim 6 is improper because that claim has been canceled. Because the claim is written in dependent form but is not dependent to the limitations of an independent claim, an artisan would be uncertain of the scope of the claimed invention. To advance prosecution and evaluate prior art, the Examiner has treated the claim 7 as depending from claim 1. This assumption seemed to reflect the subject matter of previous version of the claims.

Claim Rejections - 35 USC § 103

Basu and Thomson and Peterson

7. Claims 1-2, 4-5, 7-12, and 14-15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Basu [US Patent 6,594,629] in view of David J. Thomson, "An Overview of Multiple-Window and Quadratic-Inverse Spectrum Estimation Methods," IEEE 1994, pp. VI 185-VI 194 and Peterson et al. [US Patent 5,067,095], all already of record.

8. Regarding claim 14, Basu [at column 17] describes an apparatus for extracting visemes from an audio speech signal by describing the content and functionality of the recited limitations recognizable as a whole to one versed in the art as the following terminology:

means for receiving digitized analog speech information from the audio speech signal, means for filtering, means for converting, and means for generating [see Fig. 12, and its descriptions, especially at column 18, line 66-column 19, line 59, of the processor, memory, and software of the sampled audio (talking, speech) stream];

successive speech is frames at a fixed rate [at column 17, lines 42-43, as audio frames spaced 10 msec in time];

receiving the speech as successive speech at the fixed rate [at column 6, lines 16-18, as the every 10-msec advance of a segment of speech];

filtering each of the successive frames of digitized analog speech information to synchronously generate time domain frame vectors at the fixed rate [see Fig. 12, and its descriptions, especially at column 6, lines 14-19, of the extraction process advancing segments of sampled speech every 10 msec and extracting succeeding acoustic cepstral vectors];

wherein each of the vectors is derived from one of the successive frames [at column 6, lines 14-19, as succeeding acoustic cepstral vectors extracted from each 10-msec advance of the segment of speech];

they are classification vectors [at column 6, lines 38-40, as the probability module labeling the extracted vectors with phonemes];

synchronously generate a sequence of a set of visemes wherein each set of visemes in the sequence is derived for a corresponding one of the vectors [at column 17, lines 51-55, as assign probabilities to visemes for vectors provided to the probability module of the time instant when the audio frame occurs];

converting each frame to a spectral domain vector [at column 6, lines 20-21, as extract magnitudes of discrete Fourier transforms in a frame];

converting the spectral vectors using DCT [at column 8, lines 24-28, as transform the amplitude values, subsequently apply a discrete cosine transform].

However, Basu does not provide details of Fourier transformation to the spectral domain. In particular, Basu does not explicitly describe using prolate spheroid basis functions.

Thomson [at section 2., section 8., and section 1.] examines transformation from the time domain to the spectral domain using the discrete Fourier transform for acoustics, speech, and signal processing, and Thomson describes:

convert to a spectral domain vector using N multi-taper discrete prolate spheroid sequence basis (MTDPSSB) functions [see Eq. (18) and its description of projecting to a frequency domain by N-1 windows of a Slepian sequence (Discrete Prolate Spheroidal Wave Functions)];

they are factors of a Fredholm integral of the first kind [at page VI-186, column 1, as the projection operation of spectrum estimation is a Fredholm integral of the first kind];

N is a positive integer [see Eq. (18) and its summation limits from 0 to N-1].

As indicated, Thomson shows that using N MTDPSSB functions was known to artisans at the time of invention. Since Thomson [at page VI-188, column 1] also points out that MTDPSSB functions have the advantage of the best possible leakage properties for handling a dynamic range, it would have been obvious to one of ordinary skill in the art of converting data to the spectral domain at the time of invention to include the concepts described by Thomson at least using the MTDPSSB functions in Basu's conversion to the spectral domain because the MTDPSSB functions were known to have the advantage of the best possible leakage properties for handling a dynamic range.

Both Basu [at column 17, lines 1-17] and Thomson [at section 1.] suggest the desirability of real-time use of recognition applications, which requires the smallest possible delay between the time of occurrence of speech and the time that a recognition label for that speech is determined.

However, neither Basu nor Thomson explicitly describes a delay of latency less than 10 msec with respect to a frame with which the visemes correspond.

Like Basu, Peterson [at column 1, lines 56-59] describes speech recognition, and Peterson also describes:

latency less than 10 milliseconds [at column 11, lines 27-46, as typically 20 elements and delays of 10 microseconds (20 x .01 ms = .2 ms) to provide an output signal from the input signal].

As indicated, Peterson shows that latency less than 10 milliseconds in recognition applications was known to artisans at the time of invention. Since Peterson [at column 1, lines 44-55] also points out that neural network processing has the inherent advantage of offering real time execution, it would have been obvious to one of ordinary skill in the art of real time speech recognition at the time of invention to include the concepts described by Peterson, at least latency less than 10 milliseconds with reference to Basu's successive 10-msec frames of speech being processed for viseme recognition, because that would provide processing results as near to real time as possible within whatever certain degree of error can be tolerated.

9. Claim 1 sets forth a method with limitations comprising the functionality associated with using the apparatus recited in claim 14. Basu, Thomson, and Peterson describe and make obvious those similar limitations as indicated there; accordingly, this claim also is unpatentable.

10. Regarding claim 2, Peterson also describes:

with a latency less than 100 msec with reference to a successive frame [at column 11, lines 27-46, as typically 20 elements and delays of 10 microseconds (20 x .01 ms = .2 ms) to provide an output signal from the input signal].

11. Regarding claim 4, Basu also describes:

each set includes viseme identifiers [at column 13, lines 42-44 and 55-56, as visual speech feature vectors (visemes) labeled with phonemes];

each set includes confidence numbers [at column 13, lines 61-62, as combine with a confidence estimation that refers to a likelihood];

the confidence corresponds one to one [at column 13, lines 44-46, as each phoneme associated with visual speech feature vectors has a probability associated therewith].

12. Regarding claim 5, Basu describes the included claim elements by dependency as indicated elsewhere in this Office action. Basu also describes:

the set consists of an identity of the most likely one [at column 18, lines 5 and 53-54, as rescore the N-best list to recognize the highest likelihood];

it is a viseme [at column 17, lines 51-55, as assign probabilities to visemes].

13. Regarding claim 7, Basu, Thomson, and Peterson, describe and make obvious the included claim elements by dependency as indicated elsewhere in this Office action if the Examiner's assumption about dependency to claim 1 is correct. Thomson also describes:

multiplying a successive frame by one of the MTDPSSB functions to generate N product sets of the frame [see Eq. (18) and its description, of multiplying (for windowing) the data x_n by the N values of a Slepian sequence to generate the values of K windows];

performing a FFT of each produce set to generate N FFT sets of the frame [see Eq. (18) and its description, of the exp(-i2πf) and sum, for the FFT used for coefficient computation];

adding (change adding to combining because the addition is done to magnitude spectrums rather than separately to the real and imaginary components) together the N FFT sets of the frame to generate a summed FFT set of the frame [see page VI-187, column 1, and the example for a Simple Spectrum Estimate, of summing (combining) the square of the absolute value (magnitude) of K coefficients of the expansion coefficients from Fourier transforming].

14. Regarding claim 8, Thomson also describes:

scaling the summed FFT set of the successive frame(s) [see page VI-187, column 1, and the example for a Simple Spectrum Estimate, dividing by K the total of summing K coefficients of the expansion coefficients from Fourier transforming].

15. Regarding claim 9, Basu also describes:

a spatial classification [at column 18, lines 53-55, as rescore based on video].

16. Regarding claim 10, Peterson also describes:

by a neural network (or other) [at column 2, lines 25-28, as provide a neural network for recognizing].

17. Regarding claim 11, Peterson also describes:

a neural network [see Figs. 4a and 4b and their descriptions especially at column 6, lines 16-68, of the SPANN (sequence processing artificial neural network)];

feed-forward type [see Figs. 4a and 4b and their descriptions especially at column 6, lines 16-68, of signals-applied-at-the-inputs-processed-and-provided-through-the-outputs];

memory-less type [see Figs. 4a and 4b and their descriptions especially at column 6, lines 16-68, of signals-applied-processed-and-output];

perceptron type [see Figs. 4a and 4b and their descriptions especially at column 6, lines 16-68, of neurons].

18. Claim 12 sets forth limitations similar to claim 14. Basu, Thomson, and Peterson describe the limitations as indicated there, for a processor and software that provide the means. Basu also describes additional limitations as follows:

a processor and a memory that stores programmed instructions that control the processor [see Fig. 12, and its descriptions, especially at column 18, line 66-column 19, line 59, of the processor, memory, and software of the sampled audio (talking, speech) stream].

19. Regarding claim 15, Thomson also describes:

N is less than 5 (or other) [at page VI-193, column 2, in paragraph leading to Eq. (78), as two Slepian sequences].

Sutton and Peterson and Basu and Thomson

20. Claims 1-2, 4-5, and 7-15 are rejected under 35 U.S.C. 103(a) as being unpatentable over Sutton [US Patent 6,539,354] in view of Peterson et al. [US Patent 5,067,095], Basu [US Patent 6,594,629], and David J. Thomson, "An Overview of Multiple-Window and Quadratic-Inverse Spectrum Estimation Methods," IEEE 1994, pp. VI 185-VI 194, all already of record.

21. Regarding claim 14, Sutton [at claim 36] describes an apparatus for extracting visemes from an audio speech signal by describing the content and functionality of the recited limitations recognizable as a whole to one versed in the art as the following terminology:

means for receiving, means for filtering, means for converting, and means for generating [at column 23, lines 46-56, as computer code comprising instructions];

receiving successive frames of digitized analog speech information from the audio speech signal at a fixed rate [at column 19, lines 1-3, as receive an input stream in frames of a speech wave at a sampling rate in 10 ms frames];

filtering each of the successive frames to synchronously generate time domain vectors at the fixed rate [at column 19, lines 2-5, as compute, for each frame in 10 ms frames, a feature representation for each frame in 10 ms frames];

the vectors are frame classification vectors [at column 19, lines 5-16, as produce phoneme (phone) estimates using the window assembled from feature representations to have 16 10-ms frames, produce viseme data for the frames];

wherein each of the vectors is derived from one of the successive frames [at column 19, line 5, as compute a feature representation for each frame];

synchronously generating a sequence of a set of visemes derived from the vectors [at column 19, lines 5-16, as assemble each feature representation into a (feature) window, produce phoneme (phone) estimates using the window assembled from feature representations to have 16 10-ms frames, produce viseme data for the frames];

wherein each set of visemes in the sequence is derived from a corresponding one of the vectors [at column 26, lines 25-34, as one or more visemes active during each of the frames of a voice input is identified (for a phoneme) corresponding to each frame].

Sutton also describes:

generating them with a latency with reference to a successive frame [at column 19, lines 27-29, as the latency is around 80 ms].

Sutton [at columns 18-19] also suggests lower latency holds advantages more than the 80 ms that is explicitly described; however, Sutton does not explicitly describe latency less than 10 msec.

Like Sutton, Peterson [at column 1, lines 56-59] describes a neural network for speech recognition, and Peterson also describes:

latency less than 10 milliseconds [at column 11, lines 27-46, as typically 20 elements and delays of 10 microseconds (20 x .01 ms = .2 ms) to provide an output signal from the input signal].

As indicated, Peterson shows that latency less than 10 milliseconds was known to artisans at the time of invention. Since Peterson [at column 1, lines 44-55] also points out that neural network processing has the inherent advantage of offering real time execution, it would have been

obvious to one of ordinary skill in the art of real time speech recognition at the time of invention to include the concepts described by Peterson at least latency less than 10 milliseconds by adjusting Sutton's neural network to a latency less than 10 milliseconds with reference to the corresponding frame because that would provide less processing delay, but within whatever certain degree of error can be tolerated.

Although Sutton [at column 19, lines 1-40] describes computing a feature representation for each frame and mel-frequency cepstral coefficients (MFCC) and features, Sutton does not further discuss how to generate the MFCC frequency domain from the time domain speech. In particular, Sutton does not explicitly describe converting each frame to a spectral domain vector and converting each spectral domain vector to one of the time domain frame classification vectors.

Like Sutton, Basu [at column 17, lines 42-55] describes assigning viseme classification to a cepstral feature vector from a frame of audio speech signal. To convert the speech to classification vectors, Basu [at column 6, line 6] points out that extraction of spectral features from speech at regular intervals was know in the art of speech recognition, and Basu summarizes it as follows:

receiving the speech as successive speech at the fixed rate [at column 6, lines 16-18, as the every 10-msec advance of a segment of speech];

filtering each of the successive frames of digitized analog speech information to synchronously generate time domain frame vectors at the fixed rate [see Fig. 12, and its descriptions, especially at column 6, lines 14-19, of the extraction process advancing segments of sampled speech every 10 msec and extracting succeeding acoustic cepstral vectors];

wherein each of the vectors is derived from one of the successive frames [at column 6, lines 14-19, as succeeding acoustic cepstral vectors extracted from each 10-msec advance of the segment of speech];

they are classification vectors [at column 6, lines 38-40, as the probability module labeling the extracted vectors with phonemes];

convert each frame to a spectral domain vector [at column 6, lines 20-21, as extract magnitudes of discrete Fourier transforms in a frame];

convert the spectral vectors using DCT [at column 8, lines 24-28, as transform the amplitude values, subsequently apply a discrete cosine transform].

As indicated, Basu shows that generating cepstral features from speech was known to artisans at the time of invention. The system by Sutton requires cepstral feature, but generating them merely by any know technique from mature technologies. Sutton has not disclosed a preferred approach to those operations according to a design criterion or solution to any stated problem. Since it appears that the use of any cepstral generation that is known to artisans would perform to provide Sutton's MFCC, it would have been obvious to one of ordinary skill in the art of speech processing at the time of invention to include the concepts described by Basu, at least transforming to the spectral domain and DCT of the (logarithm) spectral features, because those steps would provide the MFCC with which Sutton's system operates.

As indicated, Sutton uses MFCC as speech features, but provides no details as to how to generate them from the speech. Basu generates the cepstral features by a conventional Fourier transform, logarithmic conversions, and discrete cosine transform (DCT). However, Basu does not provide details of Fourier transformation to the spectral domain. In particular, Basu does not explicitly describe using prolate spheroid basis functions.

Thomson [at section 2., section 8., and section 1.] examines transformation from the time domain to the spectral domain using the discrete Fourier transform for acoustics, speech, and signal processing, and Thomson describes:

convert to a spectral domain vector using N multi-taper discrete prolate spheroid sequence basis (MTDPSSB) functions [see Eq. (18) and its description of projecting to a frequency domain by N-1 windows of a Slepian sequence (Discrete Prolate Spheroidal Wave Functions)];

they are factors of a Fredholm integral of the first kind [at page VI-186, column 1, as the projection operation of spectrum estimation is a Fredholm integral of the first kind];

N is a positive integer [see Eq. (18) and its summation limits from 0 to N-1].

As indicated, Thomson shows that using N MTDPSSB functions was known to artisans at the time of invention. Since Thomson [at page VI-188, column 1] also points out that MTDPSSB functions have the advantage of the best possible leakage properties for handling a dynamic range, it would have been obvious to one of ordinary skill in the art of converting data to the spectral domain at the time of invention to include the concepts described by Thomson at least using the MTDPSSB functions in Basu's and Sutton's conversions to the spectral domain because the MTDPSSB functions were known to have the advantage of the best possible leakage properties for handling a dynamic range.

22. Claim 1 sets forth a method with limitations comprising the functionality associated with using the apparatus recited in claim 14. Sutton, Peterson, Basu, and Thomson describe and make obvious those similar limitations as indicated there; accordingly, this claim also is unpatentable.

23. Regarding claim 2, Peterson also describes:

with a latency less than 100 msec with reference to a successive frame [at column 11, lines 27-46, as typically 20 elements and delays of 10 microseconds (20 x .01 ms = .2 ms) to provide an output signal from the input signal].

24. Regarding claim 4, Basu also describes:

each set includes viseme identifiers [at column 13, lines 42-44 and 55-56, as visual speech feature vectors (visemes) labeled with phonemes];

each set includes confidence numbers [at column 13, lines 61-62, as combine with a confidence estimation that refers to a likelihood];

the confidence corresponds one to one [at column 13, lines 44-46, as each phoneme associated with visual speech feature vectors has a probability associated therewith].

25. Regarding claim 5, Basu also describes:

the set consists of an identity of the most likely one [at column 18, lines 5 and 53-54, as rescore the N-best list to recognize the highest likelihood];

it is a viseme [at column 17, lines 51-55, as assign probabilities to visemes].

26. Regarding claim 7, Sutton, Peterson, Basu and Thomson describe and make obvious the included claim elements by dependency as indicated elsewhere in this Office action if the Examiner's assumption about dependency to claim 1 is correct. Thomson also describes:

multiplying a successive frame by one of the MTDPSSB functions to generate N product sets of the frame [see Eq. (18) and its description, of multiplying (for windowing) the data x_n by the N values of a Slepian sequence to generate the values of K windows];

performing a FFT of each produce set to generate N FFT sets of the frame [see Eq. (18) and its description, of the exp(-i2πf) and sum, for the FFT used for coefficient computation];

adding (change adding to combining because the addition is done to magnitude spectrums rather than separately to the real and imaginary components) together the N FFT sets of the frame to generate a summed FFT set of the frame [see page VI-187, column 1, and the example for a Simple Spectrum Estimate, of summing (combining) the square of the absolute value (magnitude) of K coefficients of the expansion coefficients from Fourier transforming].

27. Regarding claim 8, Thomson also describes:

scaling the summed FFT set of the successive frame(s) [see page VI-187, column 1, and the example for a Simple Spectrum Estimate, dividing by K the total of summing K coefficients of the expansion coefficients from Fourier transforming].

28. Regarding claim 9, Basu also describes:

a spatial classification [at column 18, lines 53-55, as rescore based on video].

29. Regarding claim 10, Sutton also describes:

by a neural network (or other) [at column 19, lines 10-11, as include a neural network].

30. Regarding claim 11, Peterson also describes:

a neural network [see Figs. 4a and 4b and their descriptions especially at column 6, lines 16-68, of the SPANN (sequence processing artificial neural network)];

feed-forward type [see Figs. 4a and 4b and their descriptions especially at column 6, lines 16-68, of signals-applied-at-the-inputs-processed-and-provided-through-the-outputs];

memory-less type [see Figs. 4a and 4b and their descriptions especially at column 6, lines 16-68, of signals-applied-processed-and-output];

perceptron type [see Figs. 4a and 4b and their descriptions especially at column 6, lines 16-68, of neurons].

31. Claim 12 sets forth limitations similar to claim 14. Sutton, Peterson, Basu, and Thomson describe and make obvious the limitations as indicated there, for a processor and software that provide the means. Sutton also describes additional limitations as follows:

a processor and a memory that stores programmed instructions that control the processor [at column 15, lines 34-45, as a server storing all the software].

32. Claim 13 sets forth limitations similar to claim 14. Sutton, Peterson, Basu, and Thomson describe and make obvious the limitations as indicated there, for a processor and software that provide the means. Sutton also describes additional limitations as follows:

a processor and a memory that stores programmed instructions that control the processor [at column 15, lines 34-45, as a server storing all the software];

a display that displays an avatar that is formed [at column 22, lines 3-17, as a display through which a synthesis visual output according to the method has a 3D character for reading];

using the set of visemes [at column 17, lines 36-43, as viseme tracks are used to render an animation].

33. Regarding claim 15, Thomson also describes:

N is less than 5 (or other) [at page VI-193, column 2, in paragraph leading to Eq. (78), as two Slepian sequences].

Response to Arguments

34. The prior Office action, mailed July 17, 2006, objects to the claims, and rejects claims under *35 USC § 102*, citing Basu, *§ 102*, citing Sutton, and *§ 103*, citing Basu, Sutton, and others. The Applicant's arguments and changes in RESPONSE TO NON-FINAL OFFICE ACTION, filed October 17, 2006, have been fully considered with the following results.

35. With respect to objection to those claims needing clarification, the amendment removes the indicated grounds for objection. Accordingly, the objections are removed. Please see new grounds of objection.

36. With respect to rejection of claims 1, 4-5, 7-9, and 14, under *35 USC § 102* and *§ 103*, citing Basu alone and in combination, the changes entered by amendment to the independent claims include each set of visemes synchronously generated with a latency less than 10 ms with reference to its corresponding frame.

The reference Basu does not explicitly describe that limitation. Accordingly, the rejections are removed. The Applicant's assertions with respect to Basu have been considered, but they are moot in view of the new claim element. Please see new grounds of rejection applied to address the new claim element: each set of visemes synchronously generated with a latency less than 10 ms with reference to its corresponding frame.

37. With respect to rejection of claims under *35 USC § 102*, citing Sutton alone, the changes entered by amendment include each set of visemes synchronously generated with a latency less than 10 ms with reference to its corresponding frame.

The reference Sutton does not explicitly describe that limitation. Accordingly, the rejections are removed. The Applicant's assertions with respect to Sutton have been considered, but they are moot in view of the new claim element. Please see new grounds of rejection applied to address the new claim element: each set of visemes synchronously generated with a latency less than 10 ms with reference to its corresponding frame.

38. With respect to rejection of claim 11 under *35 USC § 103*, citing Sutton and Peterson in combination, the reference Sutton does not describe the whole invention of the independent claim, as amended, and the current combination of Sutton with Peterson does not make those limitations obvious compared to the prior art of record for the whole structure and interaction expressed by the combination of all limitations. Accordingly, the rejections are removed. The Applicant's assertions with respect to Sutton, Peterson, and Basu have been considered, but they are moot in view of the new claim elements. Please see new grounds of rejection.

39. With respect to rejections of claim 3 and claim 6 under *35 USC § 103*, the rejections no longer applies because the claim has been canceled.

40. With respect to rejection of claim 15 under *35 USC § 103*, citing Basu, Thomson, and Peterson in combination, and relevant to the grounds of rejection made in this Office action, the Applicant's argument appears to be as follows:

The Applicant's argument appears to be that the concepts that Peterson describes cannot be combined with the concepts of Basu or Sutton because Peterson does not describe two of the claim limitations, namely the claimed receiving and the claimed generating. This argument is not persuasive because one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. The proper approach to the issue is whether an artisan, familiar with all that Basu, Sutton, and Peterson disclose, would have found it obvious to make a solution to the problem of low-latency speech recognition corresponding to what is claimed. Peterson, when filtered through the knowledge of one skilled in the art of neural network processing, combines with Sutton and Basu to teach or suggest that latency less than 10 ms would improve on Sutton's 80 ms, neural network latency when Sutton or Basu synchronously generate visemes from one frame of digitized analog speech information.

The Applicant's remarks have been fully considered but they are not persuasive. Accordingly, the rejection of claim 15 is maintained and the rejections in the current Office action are deemed appropriate.

Conclusion

41. Any response to this action should be mailed to:

Mail Stop AF

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

or faxed to:

(571) 273-8300, (please mark "EXPEDITED PROCEDURE"; for formal communications and for informal or draft communications, additionally marked "PROPOSED" or "DRAFT")

Patent Correspondence delivered by hand or delivery services, other than the USPS, should be addressed as follows and brought to U.S. Patent and Trademark Office, Customer Service Window, **Mail Stop AF**, Randolph Building, 401 Dulany Street, Alexandria, VA 22314

42. Applicant's amendment necessitated the new ground(s) of rejection presented in this Office action. Accordingly, **THIS ACTION IS MADE FINAL.** See MPEP § 706.07(a). Applicant is reminded of the extension of time policy as set forth in 37 CFR 1.136(a).

A shortened statutory period for reply to this final action is set to expire THREE MONTHS from the mailing date of this action. In the event a first reply is filed within TWO MONTHS of the mailing date of this final action and the advisory action is not mailed until after the end of the THREE-MONTH shortened statutory period, then the shortened statutory period will expire on the date the advisory action is mailed, and any extension fee pursuant to 37 CFR 1.136(a) will be calculated from the mailing date of the advisory action. In no event, however, will the statutory period for reply expire later than SIX MONTHS from the date of this final action.

******************* IMPORTANT NOTICE *************************

The Examiner handling this application, who was assigned to Art Unit 2654, is assigned to **DIVISION 2626** as a result of consolidation in Technology Center 2600. Please include the new Division in the caption or heading of any communication. Your cooperation in this matter will assist in the timely processing of the submission and is appreciated by the Office.

43. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Donald L. Storm, of Division 2626, whose telephone number is

(571) 272-7614. The examiner can normally be reached on weekdays between 7:00 AM and 3:30 PM Eastern Time. If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Richemond Dorvil can be reached on (571) 272-7602.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Inquiries regarding the status of submissions relating to an application or questions on the Private PAIR system should be directed to the Electronic Business Center (EBC) at 866-217-9197 (toll-free) or 571-272-4100 between the hours of 6 a.m. and midnight Monday through Friday EST, or by e-mail at: ebc@uspto.gov. For general information about the PAIR system, see http://pair-direct.uspto.gov. If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

Donald L. Storm
Donald L. Storm
Examiner, Division 2626

November 14, 2006